Scoring the SAGE

Jon Cohen

American Institutes for Research

Overview

- SAGE will be entirely machine-scored
 - Immediate results will provide actionable information to students and educators
- SAGE will include challenging, machinescored constructed response items
 - Interaction Choices
 - Equations
 - "Hot Text"
 - Drawings
 - Short answer
 - Essays



How do the scoring engines work?

- Different scoring engines work differently
- Most scoring engines are Explicit
 - Scoring engine searches the student response for a match to the correct answer, answers, or characteristics of an answer supplied by the item writer
- Essay scoring engines are Implicit
 - Work by predicting how humans would score a response



Equation Scoring

- Demo—Quadratic Equation Item
- Scoring two items, rubric supplied by item writer
 - Interactions: Choose 3 points, any 3 will do
 - Equation: Compare to the Equation to the correct equation
 - Is it correct?
 - Partial credit if signs are wrong?
 - Partial credit if intercept is correct?



Hot Text Scoring

- Demo: Sherlock Holmes
- Part A is simply multiple choice, though there could be more than one right answer
- Parts B & C ask for supporting details
 - Correct answers supplied by the item writer
 - Correct answers to B&C may depend on choice in Part A
 - Item writer decides how to offer credit



Scoring Drawings

- Demo: Triangle PQR
- Graphic response rubric supplied by item writer
 - Can score the triangle itself—is it congruent?
 - Can score the location—correct translation?
 - Can score the position—correct rotation?
 - Can score the labeling



More Drawing

- Demo: Electrical Charge
- Graphic response rubric supplied by item writer
 - Can score placement of red ball
 - Can score endpoints of each arrow

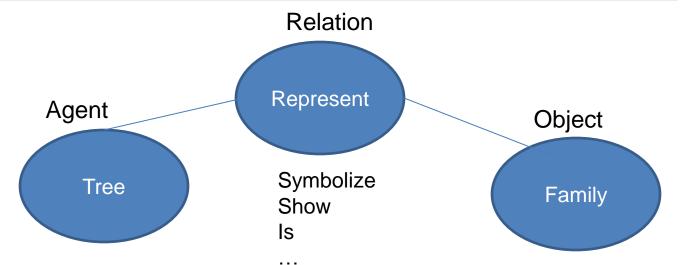


Short (Proposition) Response

- Demo: The Tree
- Item writer specifies correct "propositions"
 - Tree symbolizes
 - Family
 - Memory
 - Resilience
 - Survival



Short answer (proposition) scoring



- Concepts and relationships specified by item writer
- Pattern matching algorithm searches for evidence in student text
- Robust to spelling and grammar errors
- Works equally well for SWD, ELL, Gen Ed



Family Togetherness Loved ones

. . .

Essay Scoring

- Only implicit scoring engines
- Engine is statistically trained
 - Humans score several hundred essays
 - Measure features of the essays, including words used, spelling, grammar, etc.
 - Develop statistical model predicting humans core from measured features
- May include targeted human readbehind



Scoring Sage

- Machine scored system
- Allows real-time delivery of test results while still administering items that
 - Challenge students
 - Require students to apply skills and knowledge in realistic situations
 - Enable all students to show what they know and can do
- Diverse item types yield an exciting and authentic assessment experience



Questions?

